J. Appl. Res. Plant Sci. Vol. 4(2),702-710, 2023, https://doi.org/10.38211/joarps.2023.04.02.178

Research Article

OPEN ACCESS

Available on https://www.joarps.org
Journal of Applied Research in Plant Sciences
(JOARPS)
ISSN: 2708-3004 (Online), 2708-2997 (Print)

Genome-Wide Analysis of the Ethylene-Insensitive3-Like Gene Family in Cucumber (*Cucumis sativus*)

Rabia Mushtaq[1], Muziyen Khan[1], Mujahid Manzoor*[2], Muhammad Shafiq[1], Muhammad Bilal[3], Tariq Manzoor[5], Muhammad Ali[2], Hafiz Sabah-Ud-Din Mazhar[1], Muneeb Hashmi[1], Muazzam Anees[1], Muhammad Rizwan[4], Muhammad Saleem Haider[5]

[1]Department of Horticulture, Faculty of Agriculture sciences, University of the Punjab, Lahore, Pakistan
[2]Department of Entomology, Faculty of Agriculture sciences, University of the Punjab, Lahore, Pakistan
[3]Department of Agronomy, Faculty of Agriculture sciences, University of the Punjab, Lahore, Pakistan
[4]Department of Food Science and Technology, Faculty of Agriculture sciences, University of the Punjab, Lahore, Pakistan
[5]Department of Plant Pathology, Faculty of Agricultural Sciences, University of the Punjab, P.O BOX. 54590, Lahore, Pakistan
*Corresponding Authors Email; mujahid.iags@pu.edu.pk

Article Received 04-02-2023, Article Revised 25-03-2023, Article Accepted 04-04-2023

Abstract

The ethylene hormone identification process, which regulates the overall rate of fruit development and formation is heavily dependent on the ethylene-insensitive 3/Ethylene-insensitive3-like *(EIN3/EIL*) protein family. *EIL* harmone improve the plant's defense against both biotic and abiotic stresses. Research of the *EIL* family has been done for many plant species but in cucumber, this Gene family has not been investigated yet. Mining of the cucumber genome has identified four member of the *EIL* gene family using various bioinformatics tool. *EIL* proteins in cucumbers clustered into 4 subgroups (groups 1,2,3,4) based on the established cucumber classification. Sequence analysis and phylogeny research showed that *CsEIL3* and other *EIN3/EIL* plant proteins isolated from a progenitor signal at the time of emergence have a high degree of similarity; *CsEIN3* is involved in the flower growth process. Comprehensive genome evaluation of the *EIL* gene family in cucumber provides the ability to analyze and analyze the performance of this gene family.

Keywords; Cucumber, *EIL* protein, Genome-wide analysis, genome sequences, Domains, MEME, phylogenic tree.

Introduction

The gaseous Hormone of plant ethylene regulates signaling pathways and numerous features of plant reactions to the environment and development. An essential component in the ethylene signalling is ethylene insensitive3 *(EIN3)*. It regulates many plant reactions to growth and stress, including seed germination, cell proliferation, cell death, sex determination, fruit ripening, blossom end rot, leaf senescence, defense against pathogens and responses to mechanical stress, cold stress and salt stress(Bie *et al.*, 2013). The ethylene-insensitive protiens (EIN3/EIL) are DNA binding protein. *(EIN3 / EIL)* proteins function as important downstream transcriptional cascade regulators by binding to essential ethylene and *EIL* conserved binding sequences (ECBSs) (Yi-Qin, 2020). Analysis of gene sequences showed that *CsEIL* was different from *EILs* in other species such as soybean, wheat, and rose. While the other crops have a different number of gene sequences. For example, soybean *(Glycine max)* has 12 *EIL* gene sequences. Small homologous members *(EIL2 to EIL5*) of the *EIL* family may have minimal impact on hylene response to certain tissues and development stages or function in totally distinct ways, whereas *EIL1* plays a significant but mostly minor role in the ethylene signaling process. There are 21 *EIL* gene sequences in wheat (*Triticum astivum*). Following analysis of *EIL* 8, 17 and 21, it was discovered that these three genes regulate abiotic and biotic stress, growth and development, and phytohormone responses. *Brassica napus* have 13 Ethylene-insensitive 3 gene sequences and play a very important role in leaves of *B.napus* after hybridization and polyploidization. One of the first crops to be domesticated, cucumber (*Cucumis sativus*) is a member of the Cucurbitaceae family, which includes over 90 genera and 750 species. It is grown in practically all

nations and climate zones. Cucumber is a warm-season crop that prefers temperatures over 20 °C and is vulnerable to frost(Tatlioglu, 1993). In cucumber seedlings (*Cucumis sativus*), short-termtreatment with ethylene concentration stimulates cell division and alters cell division(Kazama, 2004) Therefore, the efficient preparation of ethylene production in different tissues at different growth stages is essential for plant development, contributing to both growth and stress growth (Pierik, Tholen, Poorter, Visser, & Voesenek, 2006). Recently, research conducted on regulating parthenocarpic fruit set in several fruit or vegetable crops, most of which are involved in hormone biosynthesis or signalling. Although parthenocarpic cucumber has been widely used in commercial production for a long time (Kaur, 2023). This work used a variety of bioinformatics methods to locate and classify genes from the EIL transcription factors in the cucumber genome. In a short, the cucumber genome was used to find EIL genes using a methodical methodology. The research was also conducted into their chromosomal location, intron/exon distribution pattern, the existence of conserved domains, and cis-regulatory elements. To determine the links between the orthologs and to explain they are likely activities, a comparative phylogenetic analysis of the EIL genes from cucumber and Arabidopsis thaliana was also conducted. Our complete genome-wide investigation of the cucumber EIL gene family members presents a reference as well as a possibility for functional analysis and also the cloning of this gene family's members in other plant species.

Materials and methods

Retrieval of sequences from databases: The protein family databse (Pfam) contained the amino acid sequence of PF04873, which is the domain that binds *EIL* DNA. This sequence was used to extract the cucumber *EIL* genes from Phytozome v13. https://phytozome-next.jgi.doe.gov/blast-search utilizing the protein-basic local alignment search technique (BLAST-P) (Finn *et al*., 2014). The NCBI CDD and SMART (Simpler Module Research Design Tool) have been used to predict the amino acid sequence (Conserved Domain Database) (https://www.ncbi.nlm.nih.gov/) using default parameters (Lu *et al.*, 2020). Genes lacking the *EIL* domain (*PF04873*) (http://pfam.xfam.org/) were removed.

Physio-chemical aspects of cucumber (*EIL*) proteins.: Using the Protparam tool (http://web.expasy.org/protparam/), the protein length (amino acid residues), molecular mass, and *EIL* protein theoretical Pi-value were acquired (Gasteiger *et al*., 2005). Phytozome was used to acquire information on IDs, chromosomal locations, peptides, CDS, and protein sequences. These *EIL* genes were given new names that mirrored their actual physical composition.

A review of gene structure: The genomic and coded sequences of the detected genes were acquired from the phytozome library to discover the intron/exon structure of *EIL*. Using the Gene Structure Display Server (GSDS) v2.0 (found at http://gsds.cbi.pku.edu.cn/), this sequence was utilized to map intron/exon. 2015 (Hu *et al*.).

Alignment of many sequences and phylogenetic analysis: Phylogenetic analysis with MEGA X v2.0 was used to align the *EIL* gene's amino acid sequence. (S. *et al*., 2018), and the tree was made using the NJ method with a 1000 bootstrap value. There were cucumber (*Cucumis sativus*), *Arabidopsis thaliana*, lettuce (*Lactuca sativa*), and sunflower (*Helianthus annus L.*) Sequences of the *EIL* protein were utilized for phylogenetic analysis. (Wang *et al.*, 2013). This analysisused to make a phylogenetic tree.

Recognition of conserved motifs and regulatory elements: The promoter encoding cis-regulatory elements in the obtained sequences were predicted using the PlantCare database. (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)(Rombauts, Déhais, Van Montagu, & Rouzé, 1999) The application Multiple EM for Motif Elicitation (MEME) was utilized to (http://meme.nbcr.net/meme/) (Bailey, Johnson, Grant, & Noble, 2015) identify number and types of motifs, the number motifs to identify was a set to 20.

Gene Duplication: Considering Ks and Ka values, the *EIL* gene family's time of split was calculated. Clustal W tool was used to align the protein sequences, and Mega X software was used to calculate the Neighbor-joining model's estimate of the Ka and Ks substitution rates. A gamma distribution with a shape parameter of 1 was used to simulate the rate variance between locations. Ka/Ks ratios were computed . According to the instructions in the software package manuals, the specifications were setting it up. To forecast the rates of molecular evolution of each pair of paralogous genes, the Ka/Ks ratios were estimated. By entering the value of Ks into the equation T = Ks /2r, where r represents a value of 6.56 10-9, the time of divergence (T) was computed.(Liu *et al.*, 2017).

Analysis of transcriptome: To assess the profile of multiple *CsEIL* genes at several plant developmental phases, the previously analyzed data of cucumber was retrieved from NCBI GEO accession (GSE151055). With the spliced reads, reads per Kilo bases per Million readable values (RPKM) from previous data and applied statistical analysis to find out the p-value and log2 fold change.

Putative micro-RNA Target Sites Analysis: To find out the mature putative micro RNA, CDS sequence was used to retrieve data from psRNA Target

https://plantgrn.noble.org/psRNATarget/analysis?function=3(Samad, 2017).

Results

Discovery of cucumber's EIL genes: The *EIN* domain sequence was utilized all cucmber sequences, sequences obtained from the Phytozome database to discover the genes of *EIL*. Four proteins were discovered by preliminary research. Proteins containing the *EIL* DNA-binding domain, were not included in the analysis. Finally, a total of 4 *CsEIL* genes were discovered and used for further analysis. (Cokol, Nair, & Rost, 2000).

Identifying conserved motifs in genes and gene structures: The exon and intron structure contributes to establishing the evolutionary link between genes or species. (Koralewski & Krutovsky, 2011). We conducted a comprehensive analysis of the exon-intron structures of cusumber *EIL was* revealed from Gene Structure Display Server (GSDS).

The MEME analysis was used to measure and map the distribution of 20 motifs over the whole *CsEIL* protein. It was observed that all 4 *CsEIL* genes showed the same motif pattern.

Phylogenetic relation with other Crops: By mapping each protein's sequence to its full length in MEGA X, a Neighbor-Joining (NJ) phylogenetic tree was created by adding boostrap values and sequances, to examine the evolutionary connections between *CsEIL* A. thaliana, *LsEIL*, and *HaEIL*. The findings demonstrated the distribution of 4 *CsEIL* proteins into 4 subgroups i.e. 1, 2, 3, and 4. The first group contains 11 *EIL* proteins of (*CsEIL2, AtEIL3, AtEIL2, LsEIL1, HaEIL7, CsEIL4, LsEIL2, HaEIL6, HaEIL8, LsEIL4, HaEIL3*). The 2nd group contains 5 *EIL* proteins of (*CsEIL3, AtEIL1, HaEIL5, LsEIL3, HaEIL1*). The 3rd group contains 3 *EIL* proteins (*CsEIL1, AtEIL4, and AtEIL6)*. The fourth group does not include any cucumber EIL gene, but consist of three EIL protiens from other species: AtEIL5, HaEIL2, and HaEIL4. (Table S1) (Fig. 1).

Table S1: Grouping of *EIL* gene family in cucumber *(Cucumis sativus), Arabidopsis thaliana, Lettuce (lactuca sativa) and sunflower (helianthus annus)* based on the phylogenetic analysis.

Group	**Cucumber (*Cucumis sativus*)**		**Arabidopsis Thaliana**		**Lettuce *(lactuca sativa)***		**Sunflower *(Helianthus annus)***	
	Gene no	Gene id	Gene no	Gene id	Gene no	Gene id	Gene no	Gene id
1	2	CsEIL2 CsEIL4	2	AtEIL2 AtEIL3	3	LsEIL1 LsEIL2 LsEIL4	4	HaEIL3 HaEIL6 HaEIL7 HaEIL8
2	1	CsEIL3	1	AtEIL1	1	LsEIL3	2	HaEIL1 HaEIL5
3	1	CsEIL1	2	AtEIL4 AtEIL6	0		0	
4	0		1	AtEIL5	0		2	HaEIL2 HaEIL4

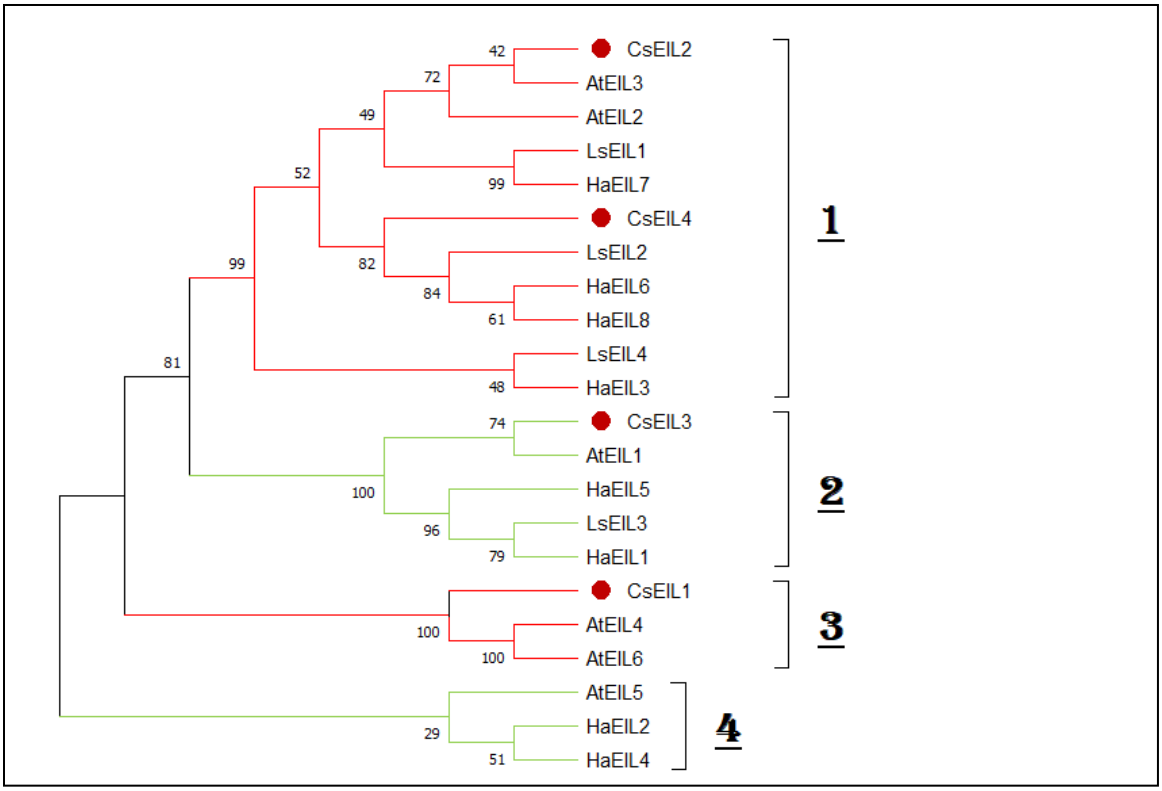

Fig. 1: Proteins where *CsEIL* proteins are marked with Red circles. The evolutionary history was inferred using the Neighbor-Joining method with 1000 Bootstrap. Evolutionary analyses were conducted in MEGA X.(Saitou and Nei 1987)(Felsenstein 1985)(ZUCKERKANDL and PAULING 1965)(Kumar *et al*. 2018) **and** Grouping of *EIL* gene family in cucumber *(Cucumis sativus), Arabidopsis thaliana, Lettuce (Lactuca sativa), and sunflower (helianthus annus)* based on the phylogenetic analysis.

Scaffold Mapping: A total of 4 *CsEIL* genes have been identified during the genome by Phytozome Database v12. *CsEIL1* was found on scaffold 397. *CsEIL2* and *CsEIL3* were found on scaffold 1079, located at a specific location on the cucumber genome. (Table 1).

EIL gene	**Source Accession**	**Chromosome**	**Chromosome Location**	**Direction**	**Size (AA)**		**pI**	**Mw**
Name	No.	No.	(bp)		mRNA length	Peptide		(KD)
CsEIL 1	Cucsa.029580	S 00397	218810..219595	F	786	262	7.10	30.2
CsEIL 2	Cucsa.142750	S 01079	958644..961672	R	1908	635	5.51	72.3
CsEIL 3	Cucsa.143220	S 01079	1540395..1542993	F	1812	603	5.39	67.7
CsEIL 4	Cucsa.364340	S 03611	1808677..1812450	R	1848	615	5.28	70.0

Table 1: Information about 4 non-redundant EIL genes discovered from the genome of cucumber Notes: AA, amino acid sequence length; MW, molecular weight; pI.

calculating non-synonymous (Ka) and synonymous (Ka) substitution rates is of great significance in reconstructing phylogeny and understanding evolutionary dynamics of protein-coding sequences across close related. This range is from 0.12 in the *CmaEIL4/CmaEIL5* pair, to 0.17 in *CmaEIL3/CmaEIL4*. The gene duplication of *CmaEIL1/CmaEIL2* is calculated to 45.07 MYA while the other 3 groups are divided from 233.40 MYA of the two pair split *CsEIL2/CsEIL4*, to 43.90 MYA for paralogous pair *CmaEIL5/CmaEIL*6. All the 3 groups' pairKa/Ks ratio

icucurbit maximum is larger than 0.3 but lower than 1 (Figure 2)

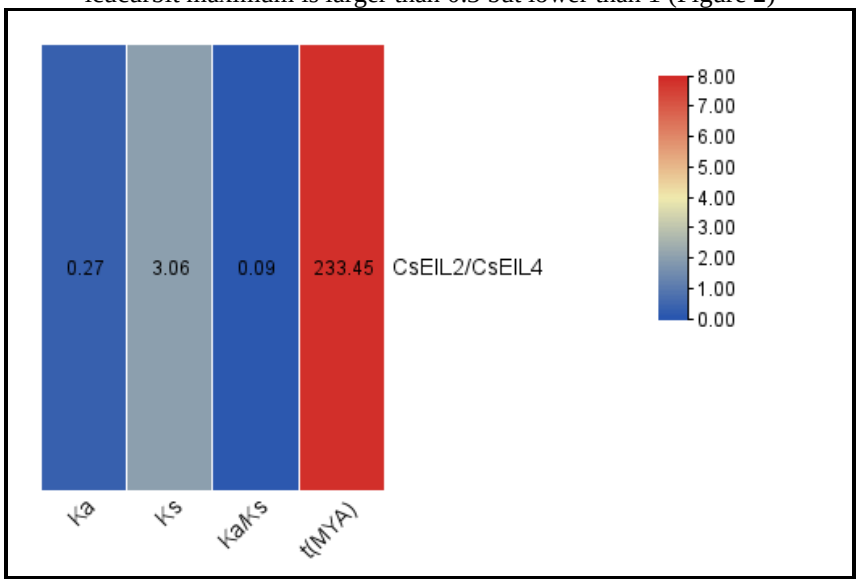

Fig. 2: Based on Ks and Ka values, the time of gene duplication was determined for several paralogous pairings of Cucumber *EIL* genes. The Nei-Gojobori model was used for the analyses. For each non-synonymous site, Ka shows the number of non-synonymoussubstitutions, while for each synonymous site, Ks shows the number of synonymous substitutions. The ratio of non-synonymous (Ka) to synonymous (Ks) mutations is represented by the expression Ka/Ks.

Analysis of Cis-regulatory elements: Physiological processes like the reaction to light are regulated by cis-regulatory components, anaerobic induction, acid response, and salicylic acid response. All 4 *CsEIL* genes contain ARE element in anaerobic induction, 2 *CsEIL* genes contain Box 4, a stored DNA module, which contributes to a simple response, 3 *CsEIL* genes possess TCA element involved in the salicylic acid response, Only one *CsEIL* gene contains the LTR element associated with low-temperature responsiveness, while 3 *CsEIL* genes displayed the wound-responsive WUN motif. The 4 *CsEIL* genes discovered cis-regulatory elements(Figure 3).

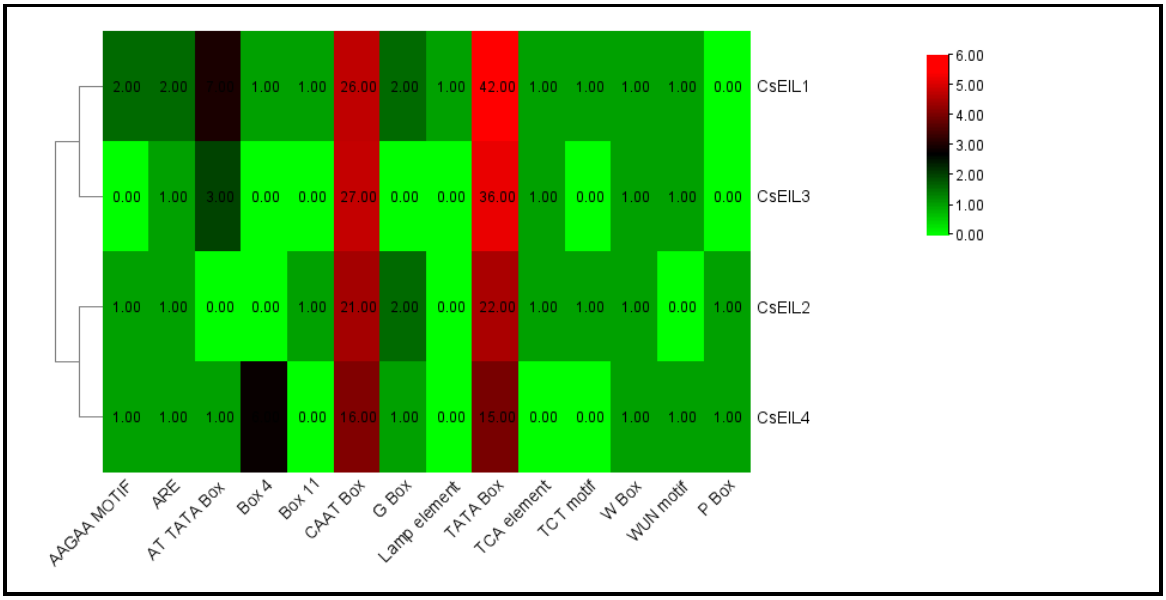

Fig. 3: Cis regulatory elements (CREs) analysis in putative *CsEIL* gene interlink with phylogenetic

Transcriptome analysis: Temperature stress : In this experiment, a cucumber variety, seedlings of '9930', was treated at 42°C temperature. The gene expression (RNA-seq) data was generated from the leaves after 0,

3, and 6 hours of high-temperature treatment stress at 42℃ (Li *et al.*, 2021)(Figure 4).

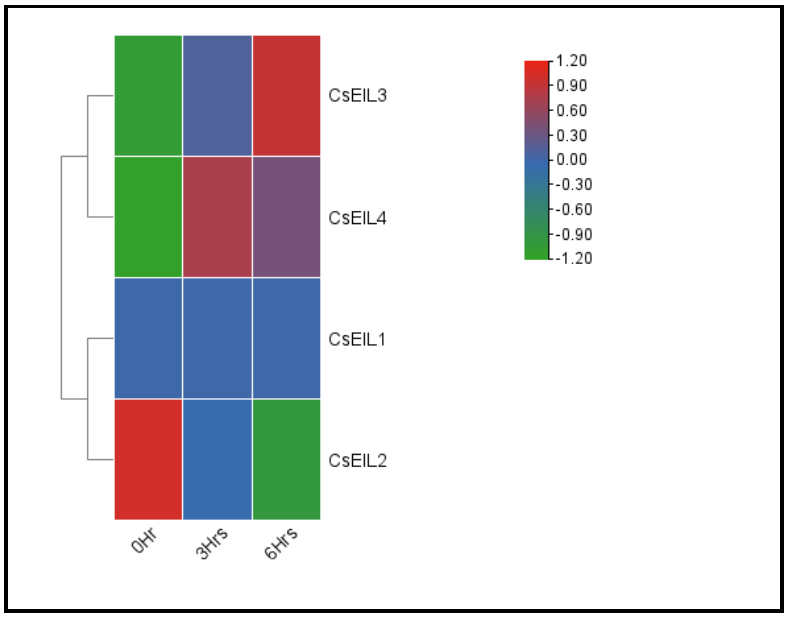

Fig. 4: Transcriptome analysis of *CsEIL*genes in cucumber by giving high-temperature stress at the different time intervals of 0, 3, and 6 hours and these genes expressed in the form of a heat map. The figure shows the different groups in different colors that colors show the expression of the gene. Red color shows more expression of gene while the green color shows minimum expression of gene (GSE151055)

Affect of Astringency on gene: This experiment was used to determine the genes that are important for the development of cucumber fruit astringency using RNA-seq.The cucumber fruit's peel and flesh were separated for RNA-seq analysis at 3, 6, and 9 days following self-pollination, generating the gene expression data(Xu *et al.*, 2019)(Figure 5).

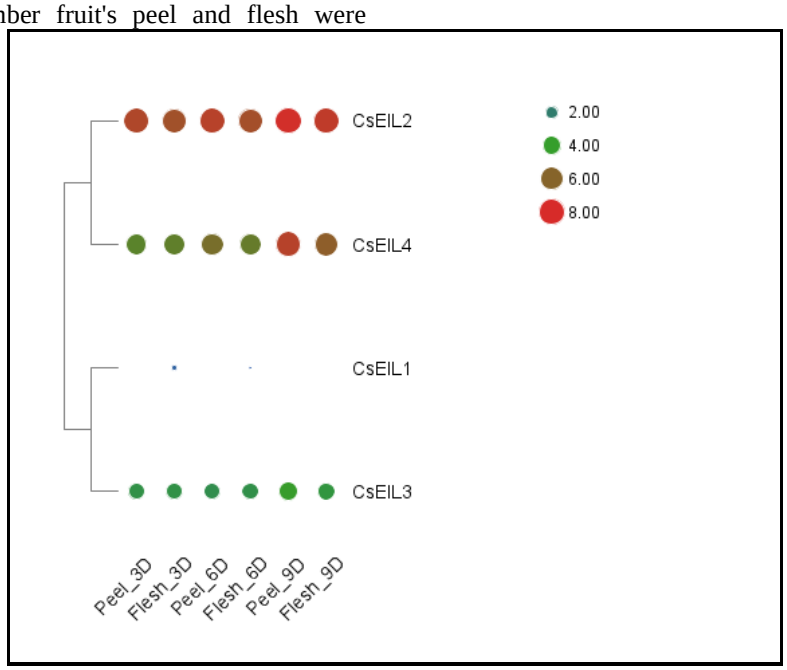

Fig. 5: Identification of candidate genes related to astringency in cucumber via RNA-seq based transcriptome profiling at different parts of cucumber like peel and flesh at different days of 3, 6, and 9 days of interval. The red circle shows more expression while the green circle shows the minimum expression of the gene and accession number is (GSE112666)

microRNA Targets: A total of 17 microRNAs were found to target 4 Cucumber *EIL* genes. *CsEIL1* gene targeted only 1 mature miRNA. *CsEIL2* targeted 4 mature miRNAs and *CsEIL3* targeted 5 mature miRNAs and *CsEIL4* targeted 7 mature miRNAs (Table S2).

Table S2: microRNA targeted prediction of *cucumber EIL*.The miRNA data was downloaded from plant *micro* RNA encyclopedia (http://pmiren.com/)

miRNA ID	Target ID	Target Start-End	Function
microRNA-122	CsEIL3	1722-1741	Not characterized yet
microRNA-79	CsEIL3	619-641	It regulates gene expression
microRNA-9	CsEIL4	793-815	Stem cell development
microRNA-44	CsEIL1	730-750	To regulate development process
microRNA-58	CsEIL2	560-580	Not characterized yet
microRNA-5667	CsEIL4	552-572	Transcription and regulation during development
microRNA-5747	CsEIL4	1201-1223	Not characterized yet
microRNA-6182	CsEIL4	1677-1699	It promote growth and development
microRNA-7499	CsEIL4	1183-1205	Not characterized yet
microRNA-7741	CsEIL3	889-909	Not characterized yet
microRNA-1	CsEIL3	1755-1775	Control the balance during differentiation
microRNA-149	CsEIL4	1738-1760	Inhibit cell migration and invasion through targeting
microRNA-153	CsEIL3	535-555	Post transcriptional regulation of plants
microRNA-287	CsEIL4	1486-1504	Plants development
microRNA-9	CsEIL2	790-812	Stem cell development
microRNA-97	CsEIL2	1765-1785	Up regulated the gene expression
microRNA-97	CsEIL2	1738-1758	Up regulated the gene expression

Discussion

Transcription factors (TFs), are regulatory molecules that take part in controlling gene expression and communication. These are protein that bind to DNA sequence to regulate gene expression. The identification and characteristics of TFs provide a better understanding of the development and growth of plants under environmental conditions.

According to the phylogenetic and domain analysis of cucumber, lettuce, and sunflower, *EIL* TFs were divided into four subfamilies (Group 1, 2, 3, 4)(Lijavetzky *et al*., 2003b), In this study, 22 *EIL* genes of cucumber, *arabidopsis thaliana* (Yang & Tuskan, 2006), lettuce, sunflower were classified into four subfamilies (Group 1,2,3,4) using the phylogenetic analysis (Figure 1).

Exon-intron structure is another amount of evidence that may be used to understand how genes or organisms have changed through time.(Bondarenko and Gelfand 2016; Koralewski and Krutovsky 2011). The predicted exon-intron association revealed that all 4 *CsEIL* genes were intron-less and on the upstream and downstream ends of the *CsEIL* domain, each was present.

The distribution of motifs among the *CsEIL* proteins reflects the phylogenetic tree's predicted evolutionary relationships between the proteins. The MEME analysis of the motif data (Gupta *et al.*, 2015; Malviya *et al.*, 2015), and domain analysis utilizing the alignment of the *CsEIL* and CDD proteins in the NCBI revealed a highly retained domain.(Lijavetzky, Carbonero, & Vicente-Carbajosa, 2003); (Gu *et al.*, 2013). Outside of the *EIL* domain, 20 different motifs have been identified that were distributed separately within the *CsEIL* genes (Figure 2). Currently, at least one or two species of nature reserves and land allocations in the *CsEIL* gene exist in the same family while there are differences between the different family members, including some similarities of *CsEIL* members within the same subfamily. In addition, *CsEIL* genetics have shown structural preservation in subsequent families and have been associated with other plants, like cucumbers, arabidopsis thaliana, lettuce, and sunflower.

In cucumber,expression of heat shock response-related *EIL* genes, the data generated from NCBI GEO and accession number (GSE151055) were retrieved. In this experiment, the seedlings of '9930', a North China-type cucumber variety, were treated at 42℃ temperature. The gene expression (RNA) data were generated from the leaves after 0, 3, and 6 hours of high-temperature treatment stress at 42℃. *CsEIL3* gene show up-regulated gene expression under high heat stress for 6 hours. While the *CsEIL2* gene shows down-regulated gene expression after 6 hours, it up-regulated gene expression after 3 hours. *CsEIL4* has minimum gene expression after 3 hours but goes down after 6 hours. *CsEIL1* shows no gene expression. Comparing the result of gene expression at different hours shows that the CsEIL4 gene is significantly expressed while the other genes do not show any significant expression.

In this study, RNA-seq was used to undertake a genome-wide analysis of gene expression in the very

astringent cucumber inbred line "YB". This data generated from NCBI GEO accession (GSE112666) were retrieved.Cucumber fruits' astringency, which harms their flavor and can give users a bad taste in their mouth, can be bothersome. Peel and flesh from the inbred line "YB" of cucumbers were separated for RNA-seq study at 3, 6, and 9 days following self-pollination. In this experiment, *CsEIL2* shows increased astringency from peel 3D to flesh 9D and *CsEIL4* shows maximum expression at peel and flesh 9D while *CsEIL1* and *CsEIL3* show no effect of astringency on any day.

Understanding the selective pressures on the displacement of amino acids are possible due to the Ka/Ks ratio. Purifying selection is also conceivable when Ka/Ks is less than 1, although the positive selection is more frequent when Ka/Ks is larger than 1. (Yang and Bielawski 2000; Hurst 2002). Analyzing selective pressure is important for understanding functional residues and functional protein modifications and frequently offers specific recommendations for amino acid sequences changed in a protein. (Morgan *et al*. 2010). The sequences from the several cucumber *EIL* groups have remarkably different Ka/Ks ratios.Despite the changes, the predicted Ka/Ks ratio is 0.089, meaning that this value is below 1, it is apparent that positive selection only altered a few sites during the evolution of the *EIL* sequences found in each group.

The most essential plant regulators, microRNA, control practically all biological functions, including growth and development, pathogen defense, and maintaining healthy internal conditions. (Carbone *et al.*, 2019; Samad, 2017; Spanudakis, 2014; Terzi, 2008). Because miRNAs are highly preserved across multiple kinds, each micro-RNA performs a particular function independent of the type to which it was addressed. When we search for detail about their functions we found that The *CsEIL*1 gene targets the miR-44 which helps the plant to regulate the development process. *CsEIL*2 targets the miR-58, miR-9, and miR97 they play different roles in plants miR-9 control stem cell development, and miR-97 regulated gene expression. *CsEIL3* gene targets the miR-122, miR-79, miR-7741, miR-1 miR-153 these miRNAs play significant roles in plants like miR-79 regulates gene expression, miR-1 control balance during differentiation, miR-153 post-transcriptional regulation of plant and it helps in the plant to regulate gene expressions. While the *CsEIL4* gene targets the miR-9, miR-5667, miR-5747, miR-6182, miR-7499, and miR-287 so the miR-9 helps plant in stem cell development, miR-5667 helps in transcription and regulation during development, miR-6182 promotes growth and development, miR-149 inhibits cell migration and invasion through targeting, and miR-287 also helps in plant development. These respective micro RNAs play a significant role in plant growth development and their function (Sun *et al*., 2015) (Figure S2).

Researchers built a phylogenetic tree with the EIL gene of other species, such as A. thaliana, lettuce, and sunflower, and separated them into three groups for the genome-wide identification of EIL in cucumber (1, 2, 3, and 4 EIL gene) . The 3 groups contained *EIL* members while the remaining 1 group has no *EIL* gene. In this study, only 1 group contained 2 *EIL* genes while the other 2 groups have one *EIL* gene (Figure **1**).

Conclusion

In this study, genetic analysis of the cucumber genome's *CsEIL* TFs have been discussed. The 4 *CsEIL* genes were split into four subgroups and the structural and functional characteristics of each *CsEIL* member were determined. The majority of *CsEIL* genes were related to fruit development and growth.Various analysis of other species has been done and also find different genes in each species like Arabidopsis thaliana has 6 EIL genes, sunflower has 8 EIL genes and lettuce has 4 EIL genes. So different species have different number of genes and functions. Also retrieve the microRNA from other article and obtained which microRNA targets which EIL gene.

Authors' Contribution

The data analysis was performed by MM, RM, MK, HSM, and MS. MS planned the experiment. This document was written by RM, MK & HS. The final manuscript was reviewed and approved by all the authors.

Acknowledgment

This work was supported by the University of the Punjab, Lahore, Pakistan.

Conflict of Interest

The authors have no potential conflict of interest.

References

Xin, T., Z. Zhang, *et al*. (2019). "Genetic Regulation of Ethylene Dosage for Cucumber Fruit Elongation."Plant Cell **31**(5): 1063-1076.

Binder, B. M. (2020). "Ethylene signaling in plants."J Biol Chem **295**(22): 7710-7725.

Bailey, Timothy L., James Johnson, Charles E. Grant, and William S. Noble. 2015. "The MEME Suite." *Nucleic Acids Research* **43**(W1): W39–49

Felsenstein, Joseph. 1985. "Confidence Limits on Phylogenies: an Approach Using The Bootstrap." *Evolution* 39(4): 783–91 (April 3, 2021).

Kumar, Sudhir *et al*. 2018. "MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms" ed. Fabia Ursula Battistuzzi. *Molecular Biology and Evolution* **35**(6): 1547–49. (April 3,

2021).

Saitou, N., and M. Nei. 1987. "The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees." *Molecular Biology and Evolution* **4**(4): 406-25. https://academic.oup.com/mbe/article/4/4/406/1029664/The-neighbourjoining-method-a-new-method-for (April 3, 2021).

Zuckerkandl, Emile, and Linus Pauling. 1965. "Evolutionary Divergence and Convergence in Proteins." In *Evolving Genes and Proteins*, Elsevier, 97–166. Higo, K., Ugawa, Y., Iwamoto, M., and Higo, H. (1998). PLACE: a database of plant cis-acting regulatory DNA elements. *Nucleic Acids Res*, 26, 358-9.

Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. The MEME suite. *Nucleic acids research*, **43**(W1), W39-W49(2015).

Bie, B., Pan, J., He, H., Yang, X., Zhao, J., & Cai, R. Molecular cloning and expression analysis of the ethylene insensitive3 (EIN3) gene in cucumber (Cucumis sativus). *Genet Mol Res*, **12**(4), 4179-4191(2013).

Carbone, F., Bruno, L., Perrotta, G., Bitonti, M. B., Muzzalupo, I., & Chiappetta, A. Identification of miRNAs involved in fruit ripening by deep sequencing of Olea europaea L. transcriptome. *PLoS One*, **14**(8), e0221460(2019).

Cokol, M., Nair, R., & Rost, B. Finding nuclear localization signals. *EMBO reports*, **1**(5), 411-415(2000).

Gu, X., Zou, Y., Su, Z., Huang, W., Zhou, Z., Arendsee, Z., & Zeng, Y. An update of DIVERGE software for functional divergence analysis of protein family. *Molecular biology and evolution*, **30**(7), 1713-1719(2013).

Gupta, S., Malviya, N., Kushwaha, H., Nasim, J., Bisht, N. C., Singh, V., & Yadav, D. Insights into structural and functional diversity of Dof (DNA binding with one finger) transcription factor. *Planta*, **241**(3), 549-562(2015).

Koralewski, T. E., & Krutovsky, K. V. Evolution of exon-intron structure and alternative splicing. *PloS one*, **6**(3), e18055(2011).

Li, C., Zhao, W., Qin, C., Yu, G., Ma, Z., Guo, Y., . Chen, J. Comparative transcriptome analysis reveals changes in gene expression in sea cucumber (Holothuria leucospilota) in response to acute temperature stress. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, **40**, 100883(2021).

Lijavetzky, D., Carbonero, P., & Vicente-Carbajosa, J. Genome-wide comparative phylogenetic analysis of the rice and Arabidopsis Dof gene families. *BMC evolutionary biology*, **3**(1), 17(2003).

Liu, H., Wu, M., Zhu, D., Pan, F., Wang, Y., & Xiang, Y. Genome-Wide analysis of the AAAP gene family in moso bamboo (Phyllostachys edulis). *BMC Plant Biol*, **17**(1), 29(2017). doi:10.1186/s12870-017-0980-z

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R.,. Marchler-Bauer, A. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res*, **48**(D1), D265-D268(2020).

Malviya, N., Gupta, S., Singh, V., Yadav, M., Bisht, N., Sarangi, B., & Yadav, D. Genome wide in silico characterization of Dof gene families of pigeonpea (Cajanus cajan (L) Millsp.). *Molecular biology reports*, **42**(2), 535-552(2015).

Pierik, R., Tholen, D., Poorter, H., Visser, E. J., & Voesenek, L. A. The Janus face of ethylene: growth inhibition and stimulation. *Trends in plant science*, **11**(4), 176-183(2006).

Rombauts, S., Déhais, P., Van Montagu, M., & Rouzé, P. PlantCARE, a plant cis-acting regulatory element database. *Nucleic acids research*, **27**(1), 295-296(1999).

Samad, A. F. A. MicroRNA and Transcription Factor: Key Players in Plant Regulatory Network. *Frontiers in Plant Science*(2017).

Spanudakis, E. The role of microRNAs in the control of flowering time. *Journal of Experimental Botany*(2014).

Tatlioglu, T. Cucumber: Cucumis sativus L *Genetic improvement of vegetable crops* (pp. 197-234): Elsevier.(1993)

Terzi, L. C. Regulation of Flowering Time by RNA Processing. *Springer Link*(2008).

Wang, G. F., Li, W. Q., Li, W. Y., Wu, G. L., Zhou, C. Y., & Chen, K. M. Characterization of Rice NADPH oxidase genes and their expression under various environmental conditions. *Int J Mol Sci*, **14**(5), 9440-

Xu, X., Pan, J., He, M., Tian, H., Qi, X., Xu, Q., & Chen, X. Transcriptome profiling reveals key genes related to astringency during cucumber fruit development. *3 Biotech*, **9**(11), 1-9(2019).

Yang, X., & Tuskan, G. A. Divergence of the Dof gene families in poplar, Arabidopsis, and rice suggests multiple modes of gene evolution after duplication. *Plant physiology*, **142**(3), 820-830(2006).